Jurnal Teknologi

Full Paper

# FREEMAN CHAIN CODE AS REPRESENTATION IN OFFLINE SIGNATURE VERIFICATION SYSTEM

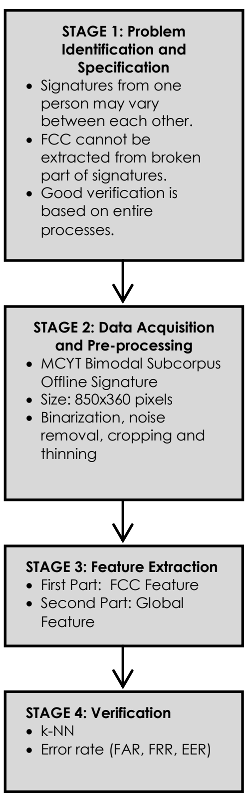
Aini Najwa Azmi[a]*, Dewi Nasien[a], Azurah Abu Samah[a]

[a]Department of Computer Science, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

**Article history**
Received
*29 November 2015*
Received in revised form
*11 April 2016*
Accepted
*17 March 2016*

*Corresponding author
anajwa2@live.utm.my

## Graphical abstract




## Abstract

Over recent years, there has been an explosive growth of interest in the pattern recognition. For example, handwritten signature is one of human biometric that can be used in many areas in terms of access control and security. However, handwritten signature is not a uniform characteristic such as fingerprint, iris or vein. It may change to several factors; mood, environment and age. Signature Verification System (SVS) is a part of pattern recognition that can be a solution for such situation. The system can be decomposed into three stages: data acquisition and preprocessing, feature extraction and verification. This paper presents techniques for SVS that uses Freeman chain code (FCC) as data representation. In the first part of feature extraction stage, the FCC was extracted by using boundary-based style on the largest contiguous part of the signature images. The extracted FCC was divided into four, eight or sixteen equal parts. In the second part of feature extraction, six global features were calculated. Finally, verification utilized k-Nearest Neighbour (k-NN) to test the performance. MCYT bimodal database was used in every stage in the system. Based on our systems, the best result achieved was False Rejection Rate (FRR) 14.67%, False Acceptance Rate (FAR) 15.83% and Equal Error Rate (EER) 0.43% with shortest computation, 7.53 seconds and 47 numbers of features.

*Keywords*: Offline signature verification system, feature extraction, Freeman Chain Code (FCC), global feature, verification

## Abstrak

Kebelakangan ini, terdapat perkembangan pesat dalam bidang *pattern recognition*. Sebagai contoh, tandatangan adalah salah satu biometrik manusia yang boleh digunakan dalam pelbagai bidang dari segi kawalan akses dan keselamatan. Walau bagaimanapun, tandatangan bukan ciri kekal seperti cap jari, iris atau urat. Ia mungkin berubah beberapa faktor; perasaan, alam sekitar dan umur. Sistem Pengesahan Tandatangan (SVS) adalah sebahagian daripada bidang *pattern recognition* yang boleh menjadi satu penyelesaian. Sistem ini boleh dikelaskan kepada tiga peringkat: pemerolehan data dan pra pemprosesan, pengekstrakan ciri dan pengesahan. Kertas kerja ini membentangkan SVS yang menggunakan teknik kod rantaian Freeman (FCC) sebagai perwakilan data. Dalam bahagian pertama peringkat pengekstrakan ciri, FCC telah diekstrak dengan menggunakan gaya sempadan berasaskan piksel sambungan yang terpanjang di dalam imej tandatangan. FCC yang diekstrak telah dibahagikan kepada empat, lapan atau enam belas bahagian yang sama. Dalam bahagian kedua pengekstrakan ciri, enam ciri-ciri global telah dikira. Akhir sekali, pengesahan menggunakan k-Nearest Neighbour (k-NN) untuk menguji prestasi eksperimen. Dataset MCYT telah digunakan dalam setiap peringkat dalam sistem. Berdasarkan sistem kami, hasil yang terbaik yang dicapai adalah False Rejection Rate (FRR) bernilai 14,6667%, False Acceptance Rate (FAR)bernilai 15,8333% dan Equal Error Rate (EER) bernilai 0,4300% dengan masa 7.534 saat dan 47 bilangan ciri.

*Kata kunci*: Sistem pengesahan tandatangan luar talian, pengekstrakan ciri, Freeman

Chain Code (FCC), ciri global, pengesah



## 1.0 INTRODUCTION

The general problem of recognising and verifying complex patterns with arbitrary orientation, location, and scale remains unsolved [1]. New and emerging applications such as data mining, web searching, retrieval of multimedia data, face recognition, and cursive handwriting recognition require robust and efficient pattern recognition techniques [1,2]. In pattern recognition, there are two problems that usually occur in an SVS. The first problem is related to the authentic user and second is related to her/his forgers. Even though these problems are not new, they are still a challenge because of the occurrence of large intra-class variations and, when forgeries are considered, small inter-class variations [3]. The importance of Signature Verification System (SVS) arises from the fact that it has long been accepted in government, legal, and commercial transactions as an acceptable method of verification [4,5,6].

The SVS has several advantages in the verification mechanism. Signature analysis can only be applied when the person is conscious and willing to write in the usual manner, although it is possible that the individual has been forced to do so [7]. But other traits like fingerprints are easier to obtain even when the person is unconscious. However, an SVS may have difficulty in discriminating between signatures since a handwritten signature is the result of a complex process, depending on the physical and psychological conditions of the signer as well as the conditions of the signing process [8, 9, 10].

A new open issue has been discussed regarding signatures across cultures [11]. Although the systems created are invariant to cultural habits and language differences, a specially designed system for different languages can perform a better verification.

SVS can be divided to two types which are offline and online system [6]. In an offline system, data acquisition is done by capturing signatures optically using a scanner and the completed signatures are available as images [12]. As a scanned signature contains a lot of noise, it must be preprocessed to produce a clean image as preparation prior to feature extraction. An online signature image usually captured by using special device can record a lot of dynamic feature such as pen speed, velocity, and pressure. Noise is very low in online image. But, in an offline system, it does not require access to special processing systems when the signatures are produced [13].

Chain code as representation is not a new method but is it still valid to use in extracting feature. It was introduced in almost five decades ago [14]. This paper proposed four techniques in pre-processing namely binarization, noise removal, thinning and cropping, followed by Freeman Chain Code (FCC) and global features in features extraction and utilized k-Nearest Neighbour (k-NN) based on Euclidean distance in verification stage. Variety of pre-processing, feature extraction and verification techniques in offline and online systems can be found in [15].

This paper is arranged as follows: current section gives an overview to what the research is all about with its process information in research framework in Section 2. In Section 3, explanation on feature extraction is detailed, continued with verification in Section 4. Next, Section 5 reveals the experimental results and comparison to previous works. Finally Section 6 draws a conclusion from result in previous section.

## 2.0 RESEARCH FRAMEWORK

To begin the research, three problems are identified to find the solutions. The first one is related to entire SVS. As signature is a type of biometric that may change with mood, environment and age, some solutions for this problem are defined. A good signature database must be updated in a few specific times so that the database is relevant to be used from time to time. Besides, a person must sign in a consistent manner to construct a series of signatures that are almost similar between each other. The second problem is related to FCC generation that failed to extract from broken parts of signature. Thus, only the largest contiguous part of the signature is chosen to extract the FCC to get the most information from the signature. Lastly, the third problem is related to verification. All processes play important role in order to achieve good verification.

The most important part in SVS is feature extraction, in which raw data representing unknown signature is transformed into effective identifier that can correctly points to its signature class, and improves classification accuracy compared to when using raw data directly. In this paper, signature raw data is defined as a sequence of Freeman chain codes (FCC), obtained by applying boundary style extraction to the image depicting a signature from a class. However, it can be done completely by comparing one-to-one process that includes data acquisition and preprocessing, feature extraction and verification.

Data acquisition is the process of sampling signals that measure real physical conditions and converting the resulting samples into digital values that can be manipulated by a computer, for example in varying colour, grey level, or binary format [16]. MCYT Bimodal Offline Signature database will be used in the entire stages. In MCYT Bimodal database, 15 genuine signatures and 15 highly-skilled forgeries with natural dynamics, which is equal to 30 signatures, were obtained for each individual so the total number of signatures in the database is 2,250 images. Forgers were given images of clients' signature to be forged and after training with them several times, they were asked to imitate the shape. Therefore, the forgeries included quick and slow imitations. All signature data were acquired with the same inking pen and same paper templates, over a similar pen tablet. The paper templates were scanned at 600 dpi [17].

In preprocessing, the image will be in binary values. Noise removal will be applied to the signature images before cropping. Finally, thinning algorithm is used to remove all redundancy by eliminating excess foreground pixels. In converting raw binary image to thinned binary image (TBI), thinning function in Image Processing Toolbox in MATLAB software is used. As the first important stage, image and data preprocessing performs the purpose of extracting regions of interest, enhancing and cleaning up the images, so that they can be directly and efficiently processed by the feature extraction component in the next stage [16].

In image processing, feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be redundant, it will be transformed into a simplified representation set of feature vector by carefully choosing relevant information from the input data. In order to perform this important task, Freeman chain code (FCC) will be used, constructed using boundary based style.

Lastly, k-NN will be used as classifier, chosen because this classifier is performing excellently in pattern recognition system [17,18]. The performance quality is measured by error rates which are False Rejection Rate (FRR), False Acceptance Rate (FAR), and Equal Error Rate (EER) in percentage.

## 3.0 FEATURE EXTRACTION

It is hard to get a perfectly match signatures from a same person. Some possibilities may occur such as variations in length, additions and deletions of portions of them, and changes in velocity due to pauses or hesitations of the writer [19]. A good database should have a series of signature from a person in order to achieve better verification. A series of signatures from a person may lead to a better feature extraction as more similar features can be extracted from more signatures.

There are two parts of feature extraction involved in this research. The first part is regarding to FCC feature. Chain code representation describes the outline for signature image by recording the direction of where is the location of the next pixel and corresponds to the neighbourhood in the image. An 8-direction FCC is used for descriptions of object borders in image field because of simplicity of the data representation and fast computation time, as shown in Figure 1. In order to extract FCC, a boundary-based style is used to minimize chain code length and it is only applied on largest contiguous part of the signature due to inability of FCC to deal with broken parts of signature [7]

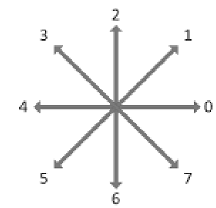



**Figure 1** 8-Neighbourhood FCC Direction

Figure 2 shows the pseudo-code of FCC extraction in boundary-based style. This paper proposed three types of chain code divisions. They are divided to four, eight and sixteen divisions for training and testing in verification stage later. For every chain code division, appearance frequency is calculated to become the directional feature. The number of features is counted with formulas in Equation 1, 2 and 3:

```
1.  Initialize data.
2.  Locate starting node (scan image from
    left to right and top to bottom to find
    the first node).
3.  Follow outermost border of the image.
4.  Stop until the tracer reaches the
    starting node again.
```

**Figure 2** Pseudo-code of applied boundary-based

$$Feature\ count = 4\ division * 8\ chain\ code = 32 \quad (1)$$

$$Feature\ count = 8\ division * 8\ chain\ code = 64 \quad (2)$$

$$Feature\ count = 16\ division * 8\ chain\ code = 32 \quad (3)$$

Next is the second part of feature extraction which is global feature. There are six global feature are calculated from the pre-processed signature images. Below are six global features that are used. Total number of global features is 15.

(1) *Signature Width*: The signature image is scanned from left to right and the distance between two points in horizontal projection is measured. This will produce one feature.

(2) *Signature Height*: The signature image is scanned from top to bottom and the distance between two points in vertical projection is measured. This will produce one feature.

(3) *Aspect Ratio*: Ratio of signature width to height. The calculation is shown in Equation 4. This will produce one feature.

$$Aspect\ Ratio = \frac{Signature\ Width, w}{Signature\ Height, h} \quad (4)$$

(4) *Distance*: The distance is measured from left to right diagonal distance of a cropped signature

image which is top right to the bottom low and top left to the bottom right. This will produce two features.

(5) *Centres of mass of all foreground pixels in an image*: It is calculated for signature image by adding all x and y locations of foreground pixels and dividing it by number of pixel counted. This will produce two features. Equations 5 and 6 are equations to find the centres of mass for x and y locations:

$$x_{centre\ of\ mass} = \sum_{x=0}^{x=lx} x\, f(x) \quad (5)$$

$$y_{centre\ of\ mass} = \sum_{y=0}^{y=ly} y\, f(y) \quad (6)$$

(6) *Counting pixel value total shift per horizontal and vertical line*: They are calculated by slicing the image horizontally into four parts and by summing shifts from black to white or white to black image. For vertical shifts, image is to be sliced vertically. This information is another unique property of signature because the chances of two signatures having exactly same shift numbers are very low. Feature count is counted as Equation 7.

$$Feature\ count = 4\ lines * (2\ directions) = 8 \quad (7)$$

## 4.0 VERIFICATION

Verification is the process of testing either a claimed signature is genuine or forgery. In our case, there are 15 signatures per class, eleven from them are trained and four are tested. Verification involved loading the template MATLAB file enrolled in the system and comparing its stored parameters. Nearest k-Neighbour (k-NN) classifier performs matching score calculation based on Euclidean distance [20]. Euclidean distance is one of the most favourite methods for measuring the distance between vectors. The performance quality is measured as False Acceptance Rate (FAR), False Rejected Rate (FRR) and Equal Error Rate (EER). The level of these two error rates depends on the decision threshold chosen, the Equal Error Rate EER is obtained when the value of them are equal (FAR = FRR) [21]. The method to obtain EER can be referred in [20]. Equations 8 and 9 show the formulas of Far and FRR.

$$FAR = \frac{Number\ of\ Falsely\ Accepted\ Images}{Total\ number\ of\ person\ in\ the\ database} \quad (8)$$

$$FRR = \frac{Number\ of\ Falsely\ Rejected\ Images}{Total\ number\ of\ person\ in\ the\ database} \quad (9)$$

## 5.0 EXPERIMENTAL RESULTS

In verification process for each signature class, a reference point is considered; if the distance between feature vector of input image and this reference point is less than a specific threshold, input image is a genuine signature, otherwise it is a forgery signature. A threshold value can be considered as a vector containing mean of corresponding elements of feature vectors in each class.

In this section, the result and some comparisons from previous work are highlighted. Table 1 shows the results from k-NN. By comparing results between chain code division of 4, 8 and 16, the best results with FRR 14.67%, FAR 15.83% and EER 0.43% are obtained from the lowest chain code division. The computation time from chain code division of four was also the shortest. However, the result obtained from chain code division of 8 and 16 do not have a huge difference. The lowest chain code division that produced the lowest number of feature gave the best results. A higher number of features could cause feature redundancy that does not help boost the system. In fact, it makes the system use up more time to execute the verification process.

**Table 1** Experimental result from k-NN

| CC Division | 4 | 8 | 16 |
|---|---|---|---|
| FRR (%) | 14.67 | 16.17 | 14.83 |
| FAR (%) | 15.83 | 16.33 | 17.50 |
| EER (%) | 0.43 | 0.45 | 0.45 |
| Computation Time (s) | 7.53 | 7.77 | 8.70 |

Figures 3 and 4 shows the trend charts for error rates which are FRR, FAR and EER and also computation time. The trends are increasing when the chain code division increases.

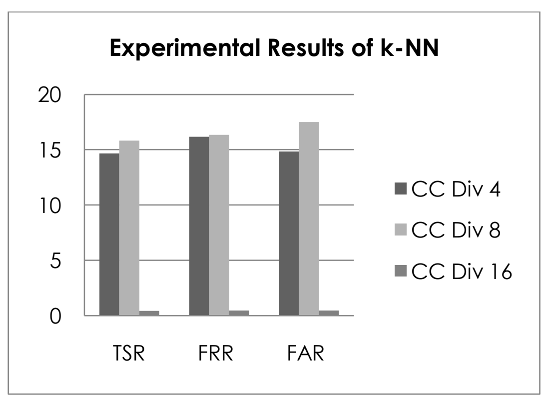



**Figure 3** Experimental Results (FRR, FAR and EER)

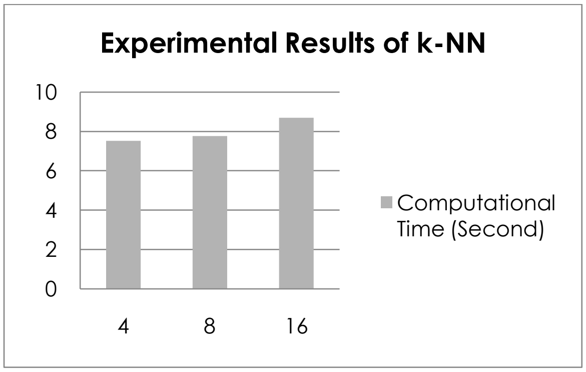



**Figure 4** Experimental Results (Computation Time)

Based on the result obtained from our system, it is satisfactory and can be competed with other systems that used the same dataset.

Comparing with work from [23], they are also used k-NN with the FAR 38.13%, FRR 38.40% and EER 22.13%. They proposed signature verification based on score level fusion of distance and orientation features of centroids. The proposed method employs symbolic representation of offline signatures using bi-interval valued feature vector.

A very good result was achieved by [4] as they were proposed pseudo-dynamic features for automatic static handwritten signature verification based on the use of grey level values from signature stroke pixels. Results have been obtained using rotation invariant uniform local binary patterns

A work from [24] that used global image level and measure the grey level variations in the image by using statistical texture features. The co-occurrence matrix and local binary pattern are analyzed and used as features.

An approach based on relative orientations of geometric centroids of split portions of signatures was proposed in [25]. The centroid orientation features of offline signatures were used to form an interval-valued symbolic feature vector for representing the signatures. They also investigated the feasibility of the proposed scheme for signature verification.

A method for verifying handwritten signatures by using Mahalanobis distance is presented in [13]. They presented two automatic measures for predicting the performance in offline signature verification. The first one measure the area of the signature slants of different directions and another one measures the intravariability of a set of signatures.

## 6.0 CONCLUSION

This paper presents an SVS that uses FCC as data representation. The raw images went through preprocessing stage which includes binarization, noise removal, cropping and thinning to produce TBI. Euclidean distance is measured and matched between nearest neighbours to find the result. MCYT Bimodal Sub-corpus database was used. Based on our systems, the best result achieved was FRR 14.6667%, FAR 15.8333% and EER 0.4300% with shortest computation, 7.534 seconds. FCC and global feature are simple and utilize less memory. There is no involvement of complicated mathematical formula and easy to understand. This is lead to shorter computation time. An optimum number of features are really important to make sure the system is working in an optimum efficiency. Based on our experiment, the increasing of chain code division will produce bigger number of features and there is no improvement for the error rate and computation time. In conclusion, the proposed SVS's results are satisfactory and can compete with recent existing systems developed by others researchers. The performance of the SVS relies very much on the information extracted from the signature image. This could be remedied in the feature extraction stage. In this study, the global features were the strong backup for the FCC features because FCC was only extracted from the largest contiguous part of the signature. As a result of the fortification, better TSR, FRR and FAR could be achieved. In addition, similar to [23], the present study used geometric centroid as a feature, but unlike theirs, this study came up with extra five global features to make sure there is enough information extracted from the signature image. This leads to better verification results.

**Table 2** Performance comparison between our work to previous works

| Author | Feature Extraction | Classification | Performance | Data Set |
|---|---|---|---|---|
| [23] | Distance and orientation feature | k-Nearest Neighbor (k-NN) | FAR: 38.13%<br>FRR: 38.40%<br>EER: 22.13% | MCYT |
| [13] | Pseudo dynamic features | Support Vector Machine (SVM) | EER: 3.42% | MCYT |
| [24] | Statistical feature | Least Squares Support Vector Machine (LS-SVM) | FAR: 9.84%<br>FRR: 13.20%<br>EER: 10.68% | MCYT |
| [25] | Symbolic representation | k- Nearest Neighbour | TSR: 60.23%<br>FRR:25.11%<br>FAR:14.66% | MCYT |
| [4] | Slant direction | Mahalanobis | FAR: 19.82% | MCYT |

| | | | | |
|---|---|---|---|---|
| | measurement | Distance | FRR: 14.85% | |
| Proposed system | FCC and global features | Euclidean distance and k-Nearest Neighbor (k-NN) | FAR: 15.83%<br>FRR: 14.67%<br>EER: 0.43% | MCYT |

## Acknowledgement

The authors are grateful to the Research Management Centre of Universiti Teknologi Malaysia (UTM), Ministry of Higher Education Malaysia for the Fundamental Research Grant Scheme (FRGS) vote number 4F264, Research University Grant (RUG) vote number 10J73 and The World Academy of Sciences-Committee on Scientific and Technological (TWAS-COMSTECH) vote number 4B197, which have facilitated the success of this project.

## References

[1] Jain, A. K., Duin, R. P., and Mao, J. 2000. Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 22(1): 4-37

[2] Basu, J. K., Bhattacharyya D., and Kim, T. 2010. Use of Artificial Neural Network in Pattern Recognition. *Journal of Software Engineering and Its Application.* 4(2): 23-34.

[3] Fierrez, J., Ortega-Garcia, J., Ramos, D., and Gonzalez-Rodriguez, J. 2007. HMM-Based On-Line Signature Verification: Feature Extraction and Signature Modeling. *Pattern Recognition Letters*. 28(16): 2325-2334.

[4] Alonso-Fernandez, F., Fairhurst, M. C., Fierrez, J., and Ortega-Garcia, J. 2007. Automatic Measures for Predicting Performance in Off-line Signature. IEEE Texas International Conference Image Processing. 1-369.

[5] Serdouk, Y., Nemmour, H., and Chibani, Y. 2016. New Off-line Handwritten Signature Verification Method based on Artificial Immune Recognition System. Expert Systems with Applications. *Elsevier.*

[6] Djeddi, C., Siddiqi, I., Al-Maadeed, S., Souici-Meslati, L., Gattal, A., and Ennaji, A. 2015. Signature Verification for Offline Skilled Forgeries Using Textural Features. *In 2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITI). IEEE*. 76-80.

[7] Vargas, V., Ferrer, M. A., Travieso, C.M., and Alonso, J. B. 2011. Off-line Signature Verification Based on Grey Level Information using Texture Features, Pattern Recognition. 44(2): 375-385.

[8] Impedovo, D., and Pirlo, G. 2011. Stability Analysis of Static Signatures for Automatic Signature Verification. In G. Maino & G. Foresti (Eds.), Image Analysis and Processing – ICIAP 2011. *Springer Berlin Heidelberg*. 241-247.

[9] Nanni, L., Lumini, A., and Brahnam, S. 2010. Local Binary Patterns Variants as Texture Descriptors for Medical Image analysis. Artificial intelligence in Medicine. 49(2): 117-125.

[10] Neamah, K., Mohamad, D., Saba, T., and Rehman, A. 2014. Discriminative Features Mining For Offline Handwritten Signature Verification. Journal of 3D Research. 5(1): 1-6

[11] Impedovo, D., Pirlo, G., and Plamondon, R. 2012. Handwritten Signature Verification: New Advancements And Open Issues. *In Conference on Frontiers in Handwriting Recognition (ICFHR),* 2012 International. IEEE. 367-372.

[12] Enqi Z., Jinxu, G., Jianbin Z., Chan M., Linjuan W. 2009. Online Handwritten Signature Verification based on Two Levels Back Propagation Neural Network. *Chengdu International Symposium on Intelligent Ubiquitous Computing and Education*. 202-205

[13] Ferrer, M., Vargas, J., Morales, A., and Ordóñez, A. 2012. Robustness of Offline Signature Verification Based on Gray Level Features. IEEE Journal of Information Forensics and Security. 7(3): 966-977.

[14] Freeman, H. 1961. Techniques for the Digital Computer Analysis of Chain Encoded Arbitrary Plane Curves. *Paper presented at the Electron.*

[15] Azmi, A. N., and Nasien, D. 2014. Freeman Chain Code (FCC) Representation in Signature Fraud Detection Based On Nearest Neighbour and Artificial Neural Network (ANN) Classifiers. *International Journal of Image Processing (IJIP).* 8(6): 434.

[16] Cheriet M., Kharma N., Liu CL., Suen C. 2007. Character Recognition Systems: A Guide for Students and Practitioners. John Wiley & Sons, Canada

[17] Pal S., Alireza A., Pal U., Blumenstein M. 2011. Offline Signature Identification using Background and Foreground Information. *Adelaide International Conference on Digital Image Computing Techniques and Applications*. 672-677

[18] Ortega-Garcia J., et al 2003. MCYT Baseline Corpus: A Bimodal Biometric Database. *IEEE Proceedings Vision, Image and Signal Processing*. 150(6): 395-401.

[19] Akram M., Qasim R., and Amin M. A. 2012. A Comparative Study Of Signature Recognition Problem Using Statistical Signatures And Artificial Neural Networks. *Dhaka International Conference on Informatics, Electronics & Vision.* 925-929,

[20] Munich, M. E., and Perona, P. 2003. Visual identification by signature tracking. *Journal of Pattern Analysis and Machine Intelligence*. 25(2): 200-217.

[21] Migual A. F., Jesus B. A., and Carlos M. T. 2005. Off-line Geometric Parameters for Automatic Signature Verification Using Fixed-Point Arithmetic. *Journal of Pattern Analysis and Machine Intelligence*. 27(6): 993-997

[22] Štruc V., Pavešic, N. 2010. The Complete Gabor-Fisher Classifier for Robust Face Recognition. *EURASIP Advances in Signal Processing.* 1-13

[23] Prakash, H. N. and Guru, D. S. 2010. Offline Signature Verification - An Approach Based on Score Level Fusion. *Journal of Computer Applications*. 1(18): 0975-8887

[24] Vargas, J. F., Ferrer, M. A., Travieso, C. M., and Alonso, J. B. 2011. Off-line Signature Verification Based on Grey Level information using Texture Features. *Journal of Pattern Recognition*. 44(2): 75-385

[25] Prakash, H. N., and Guru, D. S. 2009. Relative Orientations of Geometric Centroids for Off-Line Signature Verification. *Proceedings of the 2009 IEEE 7th International Conference Advances in Pattern Recognition*. February 4-6. Kolkata. 201-204.